DOCUMENT RESUME

ED 435 133 EA 030 117

AUTHOR Eide, Eric R.; Showalter, Mark H.
TITLE The Effect of Grade Retention on Educational and Labor Market Outcomes.
PUB DATE 1999-05-00
NOTE 35p.; Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Quebec, Canada, April 19-23, 1999).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Dropout Research; Dropouts; *Economics; Elementary Secondary Education; Expectation; *Grade Repetition; *Outcomes of Education; *School Holding Power; Student Welfare
IDENTIFIERS High School and Beyond (NCES)

ABSTRACT

This paper examines the decision-making process in retaining a student and the empirical effects of such decisions within an economic framework. The article models the decision to retain a child as a parental decision in which a parent holds a child back because the benefit of retention is higher expected earnings. The costs of retention include deferred entry into the job market and other nonpecuniary considerations. The paper contains an empirical section that uses the High School and Beyond (HSB) data set to examine the effects of retention on the probability of dropping out of high school and on labor market earnings several years after the student has entered the workforce. Information on respondents to the 1980 sophomore cohort of the HSB data set was used. The samples for white males and white females were four to five times as large as their black counterparts. The results show significant differences between males and females who are white and males and females who are black. For all demographic groups, the effect of grade retention on dropping out of high school and on labor-market earnings suggests a negative and statistically significant relationship between repeating a grade and the outcomes. (Contains 34 references and 7 tables.) (RJM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 435 133

THE EFFECT OF GRADE RETENTION ON EDUCATIONAL AND LABOR MARKET OUTCOMES

Eric R. Eide[a] and Mark H. Showalter[b]

a: Department of Economics, Brigham Young University
b: Department of Economics, Brigham Young University and
MIT Sloan School of Management

First Version: October 1998
Latest Version: May 1999

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
E. Eide
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

EA 030 117

We thank Joe Altonji, Jill Constantine, Jeff Grogger, Richard Murnane, Josh Angrist, Susan Dynarski, Eric Bettinger and participants at the MIT labor seminar for helpful comments. David Sims and Jane Anderson provided excellent research assistance.

Authors can be contacted at the following addresses: Eric Eide at eide@byu.edu; Mark Showalter at showaltr@mit.edu.

BEST COPY AVAILABLE

ABSTRACT

Grade retention is a major issue in the ongoing debate of how to improve primary and secondary education in the United States. Also known as 'flunking' or 'being held back', repeating a grade is very common. For 1995, the National Center for Education Statistics (NCES) estimated that over 13 percent of individuals ages 16 to 24 had repeated at least one grade while in school. For blacks, the estimate is nearly 1 in 5.

This paper examines the retention decision and its e.npirical effects using an economic framework. Within our model, the retention decision is endogenous with respect to such observables as dropping out of school and labor market earnings and this endogeneity needs to be accounted for in empirical work.

In the empirical section of the paper we use the High School and Beyond (HSB) data set to examine the effects of retention on the probability of dropping out of high school and on labor market earnings several years after the student has entered the workforce. This empirical work adds to the existent literature on grade school retention in several dimensions. First, the data set is nationally representative and thereby provides some generality that is lacking in much of the current literature which typically consists of case studies of individual schools or districts. Second, since the HSB is a panel of students tracked since their sophomore year in 1980 up to 1992, we are able to examine the relatively long-run impact of retention on eventual labor market earnings, which no previous study has been able to measure. Third, we account for the endogeneity of grade retention by using Instrumental Variables (IV) estimation where the key instrument is based on exogenous variation across states in kindergarten entry dates.

I. INTRODUCTION

Grade retention is a major issue in the ongoing debate over how to improve primary and secondary education in the United States. Also known as 'flunking' or 'being held back', grade retention generally refers to the educational practice of having a child repeat a grade in school. For 1995, the National Center for Education Statistics (NCES) estimated that over 13 percent of individuals ages 16 to 24 had repeated at least one grade. For blacks, the estimate is nearly 1 in 5 (NCES 97-473, Table 24, 1997).

The NCES estimates are low compared to other studies. The NCES estimates imply that roughly 1 percent of students are retained each year.[1] By contrast, a study by Rose et al. (1983) using data from the 1980s estimates that 5.5 percent of all students in grades K-12 are retained each year. With a sample of 15 states, several states had rates that implied over 50 percent of students would be retained at least once by the tenth grade.

As for more recent evidence, Table 1 gives estimates derived from the October educational supplement of the 1997 Current Population Survey (CPS). The survey asked of each child in a household what grade they were enrolled in during the current year and what grade they were enrolled in during the previous year, and we use this information to calculate retention rates. These estimates show that retention rates are substantially higher in the first grade than in subsequent grades, that blacks tend to be retained at higher rates than whites, and that males are retained more frequently than females.

---

[1] 13 percent divided by 13 grades. This is only an approximation. Multiple retakes will lower the number somewhat and as noted previously, retentions are not uniform across grade levels.

Although there are no systematic national estimates of either the number of students held back each year or of the cost to the public of this practice, a rough estimate of the direct dollar cost can be made. In 1996, the NCES estimated there were 44.7 million students in grades K-12 in the U.S. and that the average expenditure per child was almost $5,800 (NCES 97-554 1997). Using a conservative estimate of 1 percent annual retention rate, the estimated cost for retention is approximately $2.6 billion per year and affects about 450,000 children. Credible estimates of 5 percent annual retention rates imply an annual cost of approximately $13 billion and over 2 million children affected. By way of comparison, the federal Head Start program had appropriations of approximately $3.6 billion in fiscal year 1997 and served approximately 800,000 children. The retention cost estimates also obviously ignore the private costs of retention, such as the delay of entry into the labor force or matriculation into post-secondary education, which are likely to be substantial.

Despite its widespread use, retention is a very controversial practice.[2] Opponents of retention contend there is a causal link between retention and negative outcomes such as dropping out of school (Wagenaar 1987, Rumberger 1987, Fine 1991, Cairns, Cairns and Neckerman 1989, Grissom and Shepard 1989 (Chapter 3)), low self esteem, and poor academic performance generally (Kellam et al.1975, Royce, Darlington and Murray 1983). For example, Grissom and Shepard (1989) estimate that retaining a student increases the probability of not completing high school by 20 to 30 percent. Holmes (1989) surveyed 47 empirical studies using a variety of academic achievement measures. He found that retained students tended to score 0.19 to 0.31 standard deviations below comparable students who had not been retained. House

[2]See Holmes (1989) and Karweit (1992) for reviews of the research on retention.

(1989, p. 209) summarized the research as follows: "It would be difficult to find another educational practice on which the evidence is so unequivocally negative."

However, there does exist a small body of empirical evidence suggesting there are beneficial effects to retention. Most notable is the recent research by Alexander, Entwisle and Dauber (1994) which documents the educational progress of a cohort of several hundred students in the Baltimore school system over an 8 year period and finds that retention has a positive effect on academic achievement as well as on more subjective measures such as self-esteem. In other work, Kerzner (1982) and Pierson and Connell (1992) also find some positive effects on academic outcomes, although the evidence tends to be mixed.

One key element that is missing in the current debate is a serious discussion about how retention decisions are made and how to account for the decision process in measuring the effects of retention. If retention is based solely upon a student passing a well-defined academic standard, that will likely have a different set of implications than if the decision is based on some guess as to how much a given student would benefit from an additional year in a particular grade. Economists have not contributed much to this lively and important debate although the heart of the problem is a resource allocation decision.[3]

This paper examines the retention decision and its empirical effects within an economic framework. We model the decision to retain a child as a parental decision with a parent deciding to hold a child back based on the net gain to the child where the benefit of retention is higher expected earnings, and the cost includes both the deferred entry into the job market and

---

[3]The single exception we are aware of is a paper by Gomes-Neto and Hanushek (1994) which analyzes the causes and consequences of grade retention in Brazil. A few papers by economists have used retention as an outcome variable (e.g. Currie and Thomas (1995)).

nonpecuniary considerations. With such a framework, the retention decision is clearly endogenous with respect to such observables as dropping out of school and labor market earnings and this endogeneity needs to be accounted for in empirical work.

In the empirical section of the paper we use the *High School and Beyond* (HSB) data set to examine the effects of retention on the probability of dropping out of high school and on labor market earnings several years after the student has entered the workforce. This empirical work adds to the existent literature on grade school retention in several dimensions. First, the data set is nationally representative and thereby provides some generality that is lacking in much of the current literature which typically consists of case studies of individual schools or districts. Second, since the HSB is a panel of students tracked since their sophomore year in 1980 up to 1992, we are able to examine the relatively long-run impact of retention on eventual labor market earnings, which no previous study has been able to measure. Third, we account for the endogeneity of grade retention by using Instrumental Variables (IV) estimation where the key instrument is based on exogenous variation across states in kindergarten entry dates.

The outline of the paper is as follows: Section II presents the model, Section III decribes the data and estimation approach, Section IV gives empirical results, and Section V concludes.

II. Model

We begin by presenting a simple model of grade retention to illustrate some of the key assumptions underlying our analysis. The basic structure of the model is that parents decide period by period (grade by grade) whether to hold their child back or to allow the child to continue to the

following grade.[4] In our stylized framework, the decision is based on the parent's expectation of the child's labor market earnings upon completion of school: if the differential in expected earnings exceeds the cost of an additional year of schooling, then the child is retained.

That parents optimize on behalf of their child's welfare is based on Becker's model of altruistic parents (Becker, 1981). For our purposes, focusing on earnings simplifies the model substantially and leads to our empirical estimation. In a more general model, parents would optimize with respect to the child's current and expected future utility; we view earnings as a reasonable approximation to begin with.

There are 12 grades and a single period labor market following completion of the last grade.[5] There are two possible outcomes in each of the grades and in the labor market: "low" and "high." So, for example, at the end of the first grade a student is ranked either low or high on academic achievement and this pattern continues until entry into the labor market where there are two outcomes: "high earnings" and "low earnings." There exists a set of exogenous transition probabilities describing the progression from grade to grade and from the final grade to the labor market with the following properties: 1) $Pr(H_{t+1} | H_t, \alpha)$ increases at a decreasing rate over time and 2) $Pr (L_{t+1} | L_t, \alpha)$ also increases at a decreasing rate over time where Pr(x|y,α) is the probability of event x conditional on event y and α where α represents individual specific characteristics such as innate intelligence or "motivation". H_t refers to a student being ranked "high" at the end of grade t and L_t corresponds to a student being ranked "low" at the end of grade

---

[4]This model will also work if instead of parental decision, the retention decision is made by an altruistic educator. Anecdotal evidence suggests that in most districts parents do have the final say in most retention decisions.

[5]Assuming 12 grades is obviously not essential to the model.

t. The probability of a high outcome in t+1 is increasing in α and correspondingly, the probability of a low outcome in t+1 is decreasing in α.

In words, property 1 states that the conditional probability of being ranked high in grade t+1 increases over time but at a decreasing rate, where the condition is that a student is ranked high in period t and is of type α. The second assumption is similar: the conditional probability of being ranked low in period t+1 increases over time at a decreasing rate, given that the student is ranked low in period t and is of type α. Importantly, the last transition probabilities give the probability for either high or low earnings in the labor market.

This structure for the probabilities embodies the assumption that the correlation in performance between grades is weaker in the early grades than it is in the later grades; the likelihood that a high-scoring teen in the 10^{th} grade will also do well in the 11^{th} grade is higher than the likelihood of a high-scoring 1^{st} grader doing well in the 2^{nd} grade. Obviously, the correlation for low-scoring students is similarly smaller in the early grades than in the later grades.

We next assume that retaking a grade increases the probability of a "high" outcome, or in the notation given above that $Pr(H_{t+1} \mid H_t, \alpha, \text{retake grade t}) > Pr(H_{t+1} \mid H_t, \alpha, \text{no retake})$ and conversely that $Pr(L_{t+1} \mid L_t, \alpha, \text{retake grade t}) < Pr(L_{t+1} \mid L_t, \alpha, \text{no retake})$. There is also an associated opportunity cost for each year of schooling, c_t, which can be thought of as including both pecuniary and nonpecuniary costs.

In general, at any given time the optimization problem for the parents is to decide on an education strategy that maximizes expected earnings in the labor market. To simplify the analysis, we assume that the gain to retaking the most recent grade (net of the cost of retaking the grade) exceeds the net gain from any other education path which includes at least retaking one

grade.[6] This simplification reduces the problem to a decision of whether or not to retake the current grade period by period.

This relatively simple structure leads to some interesting and important implications. First, due to how the transition probabilities change over time, the longer the unbroken string of success, the higher the probability of future success. The same pattern occurs on the negative side: unbroken periods of failure tends to lead to more failure. These results illustrate the more fundamental point that in this model the probabilities are path-dependent; the probability of success in the job market depends not only upon how well the student does in the 12^{th} grade but how the student fared in previous grades. Intuitively this leads to an incentive to get on a "high" path quickly which leads to an incentive to retake grades earlier rather than later.[7]

This model also does not rule out that it might be optimal to retake a grade even for a student who gets a high outcome in a given grade. This would result if retaking raises the probability of a high outcome in the following period by a sufficient amount. A related point is that the effect of retention will vary depending upon the value of α. For example, for student A with a high value of α, retention might change the probability of a high outcome in the following grade from 80 percent to 90 percent. For student B with a relatively low value of α, the change might be from 20 percent to 40 percent.

---

[6]Such an outcome can be justified by the assumptions that: 1) the marginal change in probabilities of a high outcome decreases at the higher grades, 2) the cost for each year of schooling increases at the higher grades and 3) the marginal gain in probability from retaking declines with additional retakes.

[7]This last result can be formalized by assuming an exponential rate of decay in how much the transition probabilities change each period.

Finally, the model makes explicit that retention is endogenously determined which must be accounted for in estimation of the retention effect. Low outcomes in grade school will tend to be correlated with low earnings in the labor market. But low outcomes in grade school will also tend to trigger retention and so we would expect there to be a negative correlation between retaking a grade and labor market earnings. However a low outcome for a particular grade is not sufficient to retake; retaking will be done only if the expected increase in earnings is sufficiently large. This implies that after accounting for the endogeneity that retention should increase expected earnings.

While this model highlights the link between retention and academic and labor market outcomes, the retention decision is very likely to also be a function of other factors. Numerous studies have found that students most likely to be retained are minority males who come from poor families (Karweit 1992, Shepard and Smith 1989). Educators and parents also consider a student's emotional maturity as an important consideration in the retention decision, and hence students who are young relative to their classmates are much more likely to be held back (Brynes 1989 (Chapter 6), Karweit 1992, Shepard and Smith 1989). Such additional factors can be incorporated into the model relatively easily and in fact are helpful in econometrically identifying the effect of retention on educational and labor market outcomes.

III. DATA AND ESTIMATION

We use the 1980 Sophomore cohort of the nationally representative *High School and Beyond* data set (National Center for Education Statistics). Information on respondents, their families, and their high schools was collected in the 1980 base year survey and in the 1982 first

follow up. Subsequent follow ups (1984, 1986, and restricted 1992) collected information on the respondents' post-secondary educational and labor market experiences.

Our measure for dropping out of high school is a binary variable which equals one if the student dropped out of high school between the sophomore and senior years and equals zero otherwise. The earnings variable is the average of log earnings in 1990, 1991 and 1992, which come from the 1992 fourth follow up.[8] Using the average of log earnings reduces the effect of transitory fluctuations in earnings, thereby providing a better estimate of long run permanent earnings. The grade retention measure is a binary variable which equals one if the student reported ever having repeated a grade and equals zero otherwise. This variable is constructed from two variables in the 1982 first follow up which ask the respondent if he or she was ever held back in school before the tenth grade. One of the questions was asked to the students who dropped out of high school between the sophomore and senior years, and the other pertained to those who did not drop out.

The samples for the high school dropout models include black and white men and women who have valid data for the dropout indicator and for the retention variable. Our sample for the earnings regressions consists of black and white men who reported positive annual earnings in 1990, 1991, or 1992 and who were not enrolled in school in a year in which positive earnings

---

[8]The HSB fourth follow up did not include information about hours worked so we are not able to separate hourly wages from labor supply effects, and are therefore not able to determine part-time from full-time workers. However, Brewer, Eide, and Ehrenberg (1996) performed sensitivity analyses using comparable data from the senior cohort of the HSB and the National Longitudinal Study of the High School Class of 1972 by constructing annual earnings measures similar to those used here and comparing them to analyses controlling for hourly wages and part-time and full-time work, and found the results were qualitatively the same.

were reported.[9] Women are not included in the earnings models due to potential sample selection problems. In all models we drop observations where the instrument is missing, and exclude those who did not participate in the follow-up from which data is drawn.

The key instrument used in our estimation is the number of days (ndays) between a child's birthday in the first year that the child is eligible to enter kindergarten, and the statutorily determined kindergarten entry date which varies across states.[10] Due to the typical pattern of a uniform distribution of birthdays across days of the calender year and the presumed exogeneity of the kindergarten entry date, this instrument is likely exogenous from outcomes such as dropping out of high school and labor market earnings.[11]

The instrument, however, is plausibly related to the probability of being retained. The argument is that, holding other things equal, the younger a child is relative to potential peers (the smaller the value of ndays), the more likely he or she will be retained. If a child is relatively slow in developing socially or academically simply due to being young relative to peers, this 'slowness' might lead parents to hold the child back a year to allow for greater maturation. In effect, ndays acts as a randomly assigned variable for each student: students with low values will tend to be

---

[9]Preliminary estimation based on samples of Hispanic students has resulted in unusual patterns not consistent with those for our samples of black and white students. Alexander, Entwistle, and Dauber (1994) find similar irregularities with Hispanic samples in their study of grade retention.

[10]For example, a child born on August 31 of 1964 in a state where the entry age is 5 years old and the entry date is September 1 would have ndays=1; a child born August 15, ndays=17.

[11]To be more precise, the instrument should have no effect on the outcome variables other than through its effect on being retained.

retained with greater frequency than students with high values and we can statistically compare the outcomes of those young students who were retained with those who were not.[12]

One possible worry about such an instrument is that somehow these 'young' children will be hurt in the long run simply for being young. There is reasonably good evidence that age differences in scholastic performance at school entry are temporary. Shepard and Smith (1986, 1987) find that first grade classes typically have 8 or 9 percentile point differences in reading achievement tests between the oldest and youngest students, but that the age effect disappears by the third grade. Reynolds (1992) finds that independent of grade retention, age at school entry does not have a significant effect on a variety of Grade 4 outcomes, including mathematics achievement, an index measuring how the teacher rated the child's school adjustment, and a self-perception measure of school competence.[13]

The structure of the HSB forces some compromises in the construction of ndays. As previously stated, ndays is constructed by first calculating the number of days between the child's birthday and the cut-off date for starting kindergarten in the student's state. More specifically, we first identify the state in which the student attended high school during the sophomore year. This data is not directly available in the HSB but can be inferred from available information (see Grogger (1996) and Brewer and Ehrenberg (1994)). Next, under the assumption that the student

---

[12]This approach is similar in spirit to the work of Angrist and Krueger (1991) who use quarter of birth as an instrumental variable for education in estimating the effect of compulsory schooling on earnings.

[13]As further evidence, we regress each of our outcomes on the full set of control variables and the retention indicator, and added separately to the specification each of the instruments to find out whether relative age in the cohort affects the outcomes independent of the other variables. In all cases for each race/gender group the instruments are insignificant, suggesting the instruments only affect the dependent variables through their effect on retention.

attended kindergarten and the sophomore year of high school in the same state we calculate the difference in number of days from the student's birthday to the cut-off date for entry into kindergarten.[14] Kindergarten cut-off dates are the state legal standards for entrance into kindergarten typically at age 5 (Wolf and Kessler (1987)). Some states allow local districts to set the dates for their schools and are not included in the estimation. A table of entry dates is given in appendix Table A1.

We experimented with several variations of the instrument since we had no strong beliefs about what the functional relationship between ndays and retention ought to look like. In the estimation section we report results using three dummy variable modifications of ndays: the first dummy instrument takes a value of 1 if ndays is less than or equal to 30 and 0 otherwise; the second and third use 60 and 90 days as cutoff points, respectively. The idea is that the effect of ndays is possibly nonlinear; it might only really matter if you are within 2 or 3 months of the cutoff date and there is little effect beyond that. Likewise we also created dummy variables for each 30 day interval of ndays and include the first four of these dummy variables as a set of instruments in one reported specification.[15]

Our OLS and IV equations include individual-specific dummy variables for community residence (urban/rural), parental education, family income, and region of the country. We set all

---

[14]Data from the Bureau of the Census (1997) suggest that only about 2 percent of youth ages 5 to 17 move across state boundaries in a given year. Hence our assumption that students attend kindergarten and tenth grade in the same state is hopefully not too stringent.

[15]In an effort to avoid some of the functional form issues, we also experimented with semi-parametric estimation of basic dropout and earnings models using the methods introduced by Abadie (1998) but were ultimately unsuccessful in getting plausible estimates, probably due to the relatively small number of relevant retainees ("compliers") in the sample.

missing values of the independent variables equal to zero and include in each regression a dummy variable equal to one if the variable is missing and equal to zero otherwise. Summary statistics for the regression variables and instruments are given in appendix Table A2.

Table 2 presents descriptive statistics for some key variables in our analysis. These figures show blacks have higher retention rates and are more likely to drop out of high school than are whites, while white men have higher earnings than black men. Comparing these numbers by gender within race, we find that for both blacks and whites, men are more likely than women to be retained and to drop out of high school, although the difference in drop out rates between white men and white women is only two percentage points.

Table 3 contains differences in the outcomes between retained and non-retained individuals, shown separately by race. For each race/gender group, students who are retained are significantly more likely to drop out of high school. Black men have the largest differential between retained and non-retained students who drop out of high school, while the smallest difference is among black women. Among both black and white men, those who never repeated a grade have significantly higher earnings than those who are retained. This gap is much greater for blacks (31 percent) than for whites (15 percent).

These tabulations of the raw data suggest that those who are retained are more likely to drop out of high school and face lower earnings as adults than individuals who were never retained. While suggestive, these summary statistics may be misleading for two important reasons: they do not control for other factors, such as family background, which also may affect the outcomes under analysis, and they do not account for the possibility that there may be some

unobserved attribute of the individuals which would affect both the outcome (e.g. earnings) and the likelihood of being held back a grade. We take account of these issues in the next section.

To test whether grade retention affects our outcomes, we consider the following estimation equation:

$$(1) \qquad Y_i = X_i\beta + R_i\delta + \varepsilon_i$$

The dependent variable Y_i represents a given outcome, either dropping out of high school or earnings, for the *i*th individual. The vector X_i contains a set of individual-specific characteristics, R_i is an indicator variable which equals one if the *i*th individual repeated a grade and equals zero otherwise, ε_i is a zero mean error term, and β and δ are parameters to be estimated. In particular, δ provides an estimate of the effect of grade retention on the outcome. The endogeneity problem arises due to the correlation of R_i and the error term, and we account for this by using Instrumental Variables estimation. Our high school dropout variable is binary, so estimating (1) by OLS amounts to estimating a linear probability model for these outcomes. We estimate both the high school dropout and earnings models by linear IV. In the high school dropout models, we rely on findings by Angrist (1991) who, using Monte Carlo methods,

demonstrated that in qualitative response models with endogenous regressors linear IV estimators perform about the same as correctly specified maximum likelihood estimators.[16]

V. RESULTS

Table 4 presents the first set of regression results which focus on the effect of retention on dropping out of high school. In overview it is useful to note that the samples for both white males and white females are 4 to 5 times as large as the black counterparts: 3046 for white males, 3291 for white females, 708 for black males, and 764 for black females. Also included in the table is the number of sampled schools: 286 and 288 for black males and females respectively; 697 and 699 for white males and females respectively.

Column 3 gives the estimated coefficient on retention using an OLS specification by gender and race classification. For all groups the estimated coefficient is positive and statistically significant with the estimates ranging from a low of 0.051 for black females up to 0.137 for black males. Interpreted casually, the estimate for white males, for example, would imply that retaining a child results in an 8.8 percent increase in the probability of dropping out of high school. While these estimates are lower than the simple tabulations in Table 3, they do suggest a strong effect of retention even after controlling for other observable characteristics.

---

[16]We chose to use linear IV over maximum likelihood estimation for a few reasons. In qualitative response models with endogenous regressors, linear IV is not only computationally more straight forward, but is also conceptually clearer in the sense that the effects of the identifying variable are more readily apparent (Angrist (1991) outlines a number of potential advantages to linear IV over maximum likelihood in this setting). Further, since we are also examining annual earnings (a continuous variable) as an outcome, linear IV gives us a single estimating framework for both dependent variables.

However, the model outlined in Section II suggests that OLS is not the right estimation approach since the retention variable will likely be endogenous. To account for this problem, Columns 4 through 8 summarize the results of several IV specifications on the various demographic groups. Column 4 gives the results using the ndays variable discussed in Section III. Columns 5 through 7 use dummy variable transformations of ndays: the instrument in Column 5 is a dummy variable which takes a value of 1 if ndays is (weakly) between 0 and 30 and 0 otherwise; Columns 6 and 7 use dummies which have 0 to 60 and 0 to 90 as the cutoffs, respectively. Column 8 uses a set of 4 dummies as instruments: 1) 1 if 0<=ndays<30, 2) 1 if 30<=ndays<60, 3) 1 if 60<=ndays<90 and 4) 1 if 90<=ndays<120.[17]

The IV results suggest some interesting and somewhat disturbing patterns. First note the "first stage" results for the four demographic groups. This is the regression of the retention variable on the instrument(s) and the other control variables, and the reported coefficient is the coefficient on the instrument along with its estimated standard error in parenthesis (or the p-value for the F-test in the case of Column 8).[18] For white males and females, the estimates are all of the expected sign and statistically significant at the 5 percent level. For example, the estimated coefficient on ndays is -0.000164 which is interpreted to mean that a person whose birthdate is 30 days from the cutoff date will be about 0.5 percent less likely to be retained than a similar student whose birthday falls exactly on the cutoff date (30*-0.000164).

---

[17]We experimented with a number of other specifications, but in general the results suggested a roughly linear effect in the 2 or 3 months prior to the cutoff date and little effect thereafter.

[18]All standard errors are robust to general heteroskedasticity and corrected for sample clustering by school. See Deaton (1997) for the effect of clustering on estimated standard errors.

Still looking at white males, the dummy variable instruments suggest much stronger effects. The estimated coefficients for the three regressions are 0.060, 0.053 and 0.044 where, for example, the 0.060 has the interpretation that a student whose birthday is within 30 days of the cutoff is 6.0 percent more likely to be retained than a student whose birthday falls beyond that range. Unsurprisingly the coefficients monotonically decline the larger the window of time included by the dummy variable. The likely explanation for the difference in implied effect of retention between the dummy variable estimates and the ndays estimate is that ndays gives equal weight to days far beyond the cutoffs used by the dummy estimates and that doesn't seem to be supported in the data. For white females the pattern is very similar both in terms of the size of the estimated coefficients, their statistical significance and the pattern revealed across the different regressions. In general, for both white males and white females the first stage results suggest that ndays or its permutations work as reasonably strong instruments for retention.

However, for black males and females the story is quite different. Only one of the ten first stage regressions results in significant coefficients and the patterns evident with whites are not nearly as apparent with these demographic groups. The difference might be simply a matter of sample size: sample sizes are much smaller for both black males and black females than for their white counterparts and this will affect the precision of the estimated coefficients.

As one check on this explanation, we took a random sample from the white groups to more closely approximate the sample sizes for blacks both in terms of the number of students and in terms of the number of schools from which students were sampled. In summary, the first stage results looked substantially weaker with this resampling, but the pattern across the results looked similar to the full sample for whites. For example, for white males the estimated coefficients in

the first stage regression on the three dummy variables for being young were 0.042 (0-30 days), 0.010 (0-60 days) and 0.007 (0-90 days), and for ndays the estimated coefficient was -0.0000307; but in no case was the coefficient estimate statistically different from zero at the 10 percent level.

Below we report another test related to sample size where we combine the HSB with data from the National Education Longitudinal Study of 1988 (NELS:88) in order to increase our sample size. Combined with the results from that exercise, there is some evidence that perhaps sample size is driving the difference between blacks and whites; but it is less than compelling. In the end we are uncertain what causes the difference and due to the weakness of the first stage results the full IV estimates should be considered fatally flawed for blacks. Nonetheless, we present the IV estimates for black in the interest of symmetry of presentation and to provide a contrast with the results for white males and white females.

The IV results for white males and white females hint at very strong positive effects of retention but in the end are ultimately inconclusive. For white males the effect of retention results in a decrease in the probability of dropping out of between 13.9 percent and 21.2 percent, depending on the instrument. The coefficient sign change from the OLS results is consistent with the theoretical model, however in no case is the effect statistically significant. The results for white females is very similar: in all cases the coefficient sign switches to negative indicating that retention lowers the probability of dropping out, but again the results are insignificant.

Given the results from the first stage regressions, it seems unlikely that the insignificance is being driven by weak instruments. However, using the interpretation of Angrist, Imbens and Rubin (1996), it is possible that the instrument is picking up a sufficiently narrow group of individuals that it makes statistical insignificance unsurprising. Under the Angrist et al

interpretation, the binary instruments act as an exogenous "treatment" or experiment, and so the regression is essentially mimicking an experiment of taking a group of students born within a few months of the cutoff date, randomly holding some of them back and comparing the outcomes between these two groups. Such an interpretation leaves open the possibility that the consistently estimated effects would be statistically significant given a large enough sample size.

Table 5 gives the results using the average of log earnings as the outcome. Only results for males are reported to avoid dealing with the issues of sample selection which tend to be much more important with female labor supply. The results are very similar in spirit to the drop out regressions. To begin with, the OLS results for both groups show a negative effect of retention on log earnings: for white males the estimated coefficient is -0.092 with a t-statistic of 2.4 while for black males it is -0.152 with a t-statistics of 1.4.

The IV regressions also look similar to the drop out results. The instruments work well for white males and poorly for black males. In accordance with the model, estimated effects of retention on earnings is estimated to be uniformly positive for white males with estimated coefficients ranging from 0.178 to 0.316, but all are imprecisely estimated. The IV results for black males are completely implausible and reflective of the poor results of the first stage regressions.[19]

---

[19]In our earnings models we estimated specifications with and without controls for educational attainment. The rationale for including educational attainment in the regression is that retention is likely highly correlated with educational attainment, and so if we omit education from the model then part of the estimated retention coefficient will reflect differences in completed schooling, rather than the retention effect per se. The potential problem with including educational attainment in the model is that it may be endogenous; therefore we estimated the models with and without education controls to provide a range of estimates and found the results were not sensitive to the inclusion of educational attainment.

To conclude the results section of the paper we briefly report findings from some additional tests we conducted to check the robustness of our results. As noted previously, one possible reason for the lack of statistical significance in the IV regressions is that the sample size is too small. To increase the sample size, we pool our HSB data with data from the National Education Longitudinal Study of 1988 (NELS:88), conducted by the National Center for Education Statistics. The NELS:88 is a nationally representative sample of 8th grade students in 1988, and is designed to be comparable to the HSB. We use the 1990 first followup of the NELS:88 to construct a sample of students who are sophomores in 1990 which we can pool with the HSB sample and approximately double our sample size. Since the most recent NELS:88 survey was conducted in 1992 when the NELS:88 respondents were high school seniors, we do not have earnings data for these individuals and we focus solely on the models for dropping out of high school. One major drawback of the NELS:88 for our purposes is that birth day is not included and so our measure of ndays and its counterparts can not be exactly duplicated thus adding in some measurement error that is not present in the HSB-only sample.

Since the HSB sample is a representative sample of sophomores, whereas the NELS:88 cohort was first interviewed in the 8th grade, we needed to make a few sample restrictions on the NELS:88 before pooling the two datasets. Specifically, we deleted from the sample anyone who was not a sophomore in 1990. Most of the variables in the HSB and the NELS:88 are directly comparable, with the exception of the family income variables. The income categories are defined differently between the two surveys, and to account for this we created mutually exclusive family income dummy variables for each cohort, and include all of the dummies in the

regressions. We also created a cohort-specific dummy variable which equals one if the respondent is from the HSB and equals zero if the respondent is from the NELS:88.

In general, the first stage results from this exercise are stronger than the results based just on the HSB. Most noticeably, the first stage results for the samples of black men and women were statistically significant at at least the 10 percent level in many cases. However, the second stage results of the high school dropout models for the pooled sample are not improved from the second stage results for the HSB sample. Specifically, the IV coefficients for the retention variable are all statistically insignificant at any conventionally acceptable level, largely reflecting what is found with the HSB-only sample.

V. CONCLUSION

Grade retention is a commonly used, yet controversial practice. A common purpose of grade retention is to provide students an additional year in which to catch up either academically or emotionally to the level of their peers; however, many researchers have found that repeating a grade in school negatively affects a student's educational career in a variety of ways. In this paper we examine the retention decision and its empirical effects within an economic framework, paying particular attention to the potential endogeneity of the retention decision.

We find significant difference in IV estimation results between males and females who are white versus males and females who are black. We are unable to determine conclusively whether this difference is attributable to a smaller sample size for blacks or whether the underlying decision process is different.

For all demographic groups, the OLS estimates of the effect of grade retention on dropping out of high school and on labor market earnings suggest a negative and statistically significant relationship between repeating a grade and the outcomes. These findings are consistent with much of the previous empirical research on grade retention. However, the IV estimates differ markedly from the OLS estimates. For whites, the estimated IV coefficients are opposite in sign from the OLS coefficients which suggests that grade retention may have some benefit to students by both lowering drop out rates and raising labor market earnings. We cannot strongly state, however, that grade retention is a useful policy since the IV estimates tended to be statistically indistinguishable from zero, although it is important to note that the parameter estimates hinted at relatively large effects. For blacks, the IV approach gave very poor first stage results and the resulting IV estimates are highly suspect.

These findings represent the gross effect of grade retention on the probability of dropping out of high school and on earnings, and do not reflect the monetary costs associated with financing an additional year of schooling for a child who repeats, nor do they capture the forgone earnings associated with the individual's delayed entry into the workforce. Given such costs, our findings on the possible benefits of retention should be interpreted cautiously.

REFERENCES

Alexander, K.L, D.R. Entwisle, S.L. Dauber (1994), On the Success of Failure, New York: University of Cambridge Press.

Angrist, J.D. (1991), "Instrumental Variables Estimation of Average Treatment Effects in Econometrics and Epidemiology." NBER Technical Working Paper No. 5072.

Angrist, J.D., G.W. Imbens, and D.B. Rubin (1996), "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, pp. 444-455.

Angrist, J.D. and A. Kreuger (1991), "Does Compulsory School Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 106, 979-1014

Becker, G.S. (1981), A Treatise on the Family, Cambridge:Harvard University Press.

Brewer, D.J., E. Eide, and R.G. Ehrenberg (1996), "Does it Pay to Attend an Elite Private College? Cross Cohort Evidence on the Effects of College Quality on Earnings," National Bureau of Economic Research Working Paper 5613.

Bound, J., D.A. Jaeger, and R.M. Baker (1996), "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, pp. 443-450.

Byrnes, D. and K. Yamamoto (1986), "Views on Grade Repetition," *Journal of Research and Development in Education*, 20, pp. 14-20.

Cairns, R.B., B.D. Cairns, and H.J. Neckerman. (1989) "Early School Dropout: Configurations and Determinants," *Child Development,* 60, pp.1437-52.

Currie, J. and D. Thomas (1995), "Does Head Start Make a Difference?" *American Economic Review*, 85, pp. 341-364.

Deaton, Angus (1997), The Analysis of Household Surveys, Baltimore, Maryland: The John Hopkins University Press.

Ehrenberg, R.G. and D.J. Brewer (1994), "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond," *Economics of Education Review*, 13, pp. 1-17.

Fine, M. (1991) Framing Dropouts: Notes on the Politics of an Urban Public High School, Albany NY: SUNY Press.

Gomes-Neto, J.B. and E.A. Hanushek (1994), "Causes and Consequences of Grade Repetition: Evidence from Brazil," *Economic Development and Cultural Change*, 43, pp. 117-148.

Grissom, J.B., and L.A. Shepard (1989) "Repeating and Dropping Out of School," in Flunking Grades: Research and Policies on Retention, Edited by L.A. Shepard and M.L. Smith, pp. 34-63. London: Falmer.

Grogger, J. (1996), "Does School Quality Explain the Recent Black/White Wage Trend?" *Journal of Labor Economics*, 14, pp. 231-253.

Holmes, C.T. (1989) "Grade Level Retention Effects: A Meta-analysis of Research Studies," in Flunking Grades: Research and Policies on Retention, Edited by L.A. Shepard and M.L. Smith, pp. 16-33. London: Falmer.

House, E.R. (1989) "Policy Implications of Retention Research," in Flunking Grades: Research and Policies on Retention, Edited by L.A. Shepard and M.L. Smith, pp. 202-213. London: Falmer.

Karweit, N.L. (1992) "Retention Policy," In Encyclopedia of Educational Research, Edited by M. Alkin, pp. 1114-1118, New York: Macmillan.

Kellam, S.G., J.D. Branch, K.C. Agrawal, and M.E. Ensminger (1975), Mental Health and Going to School: The Woodlawn Program of Assessment, Early Interyention, and Evaluation. Chicago: University of Chicago Press.

Kerzner, R.L. (1982), "The Effect of Retention on Achievement," Union, NJ: Kean College of New Jersey.

National Center for Educational Statistics, Statistical Analysis Report 97-473, July 1997.

National Center for Educational Statistics, Statistical Analysis Report 97-554, July 1997.

Pierson, L.H., and J.P. Connell (1992), "Effect of Grade Retention on Self-System Processes, School Engagement, and Academic Performance," *Journal of Educational Psychology*, 84, pp. 300-307.

Reynolds, Arthur J. (1992), "Grade Retention and School Adjustment: An Exploratory Analysis," *Educational Evaluation and Policy Analysis*, 14, 2, pp. 101-121.

Rose, J.S., Medway, F.J., Cantrell, V.L. and Marus, S.H. (1983), "A Fresh Look at the Retention-promotion Controversy," *Journal of School Psychology*, 21, pp. 201-11.

Royce, J.M., R.B. Darlington, and H.W. Murray (1983), "Pooled Analysis: FindingsAcross Studies," in As the Twig is Bent: Lasting Effects of Preschool Programs, Edited by the Consortium for Longitudinal Studies, pp. 411-459, Hillsdale, NJ: Erlbaum.

Rumberger, R.W. (1987) "High School Dropouts: A Review of Issues and Evidence," *Review of Educational Research*, 57, pp. 101-121.

Shepard, L.A., and M.L. Smith (1986), "Synthesis of Research on School Readiness and Kindergarten Retention," *Educational Leadership*, 43, pp. 78-86.

Shepard, L.A., and M.L. Smith (1987), "Effects of Kindergarten Retention at the End of First Grade," *Psychology in the Schools*, 24, pp. 346-357.

Shepard, L.A., and M.L. Smith (1989), "Introduction and Overview," in Flunking Grades: Research and Policies on Retention, Edited by L.A. Shepard and M.L. Smith, pp. 1-15. London: Falmer.

Smith, M.L. (1989) "Teacher's Beliefs about Retention," in Flunking Grades: Research and Policies on Retention, Edited by L.A. Shepard and M.L. Smith, pp. 132-150. London: Falmer.

Wagenaar, T.C. (1987) "What Do We Know About Dropping Out of High School?" In Research in Sociology of Education and Socialization, Edited by A.C. Kerckhoff, pp. 161-90, New York: Russell Sage.

Wolf, J.M, and A.L. Kessler (1987) "Entrance to Kindergarten: What is the Best Age?", ERS Monograph, Educational Research Service, Arlington Virginia.

Table 1
Estimated Retention Rates (Ages 14 and Under)

Grade	All	Whites	Blacks	Males	Females
1	6.69	5.70	14.59	7.81	5.54
2	2.22	2.26	1.86	2.87	1.48
3	1.57	1.20	3.32	1.78	1.37
4	1.62	1.52	2.42	1.68	1.56
5	1.23	0.99	2.07	1.42	1.01
6	1.36	1.03	2.66	1.71	1.03
7	1.80	1.89	1.28	2.49	1.08
8	1.10	0.99	2.02	1.60	0.58
Average	2.24%	1.97%	3.98%	2.71%	1.74%

Source: Current Population Survey, Oct 1997, authors' computations

Table 2

Sample Means by Gender and Race

	Men		Women	
Variable	**White**	**Black**	**White**	**Black**
Percent Retained	15.80	21.00	9.80	16.90
	(36.5)	(40.8)	(29.8)	(37.5)
Percent Dropped Out of High School	16.80	22.00	15.00	16.10
	(37.4)	(41.5)	(35.7)	(36.8)
Average Log Earnings	9.39	9.15	---	---
	(0.69)	(0.89)		

Note: Estimated standard deviations in parentheses. Sample sizes are as follows: Black Men, 708; Black Women, 764; White Men, 3046; White Women, 3291. For average earnings: Black Men, 477; White Men, 2551. Average Log Earnings are average of 1990-1992 earnings in 1984 constant dollars.

Table 3

Differences in Average Outcomes of Retained and Non-Retained Students, by Gender and Race

	Men		Women	
	White	**Black**	**White**	**Black**
Dropping Out of High School				
(1) % of retained students who drop out of high school	38.0 (2.2)	45.0 (4.1)	35.0 (2.7)	27.1 (3.9)
(2) % of non-retained students who drop out of high school	12.9 (0.7)	15.9 (1.5)	12.8 (0.6)	13.9 (1.4)
(3) Difference	25.1 (2.3)	29.1 (4.4)	22.2 (2.8)	13.2 (4.1)
Average Log Earnings				
(1) If ever retained	9.27 (0.04)	8.91 (0.10)	---	---
(2) If never retained	9.42 (0.01)	9.22 (0.04)	---	---
(3) Difference	-0.15 (0.04)	-0.31 (0.11)	---	---

Note: Standard errors (for the estimated mean) in parentheses. Average log earnings are for 1990-1992 in 1984 $'s.

Table 4

Estimates of the Effect of Grade Retention on Dropping Out of High School by Gender and Race

			OLS Estimates	Instrumental Variable Estimates				
					Binary Measures			
	Sample Size	*# of Schools*		*ndays*	*ndays<=30*	*ndays<=60*	*ndays<=90*	*Month Indicators*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
White Males	3046	697						
1st Stage				-0.000164**	0.060**	0.053**	0.044**	
				(0.000063)	(0.026)	(0.019)	(0.016)	0.04
Estimated Coefficients			0.088**	-0.139	-0.177	-0.141	-0.209	-0.212
			(0.016)	(0.276)	(0.287)	(0.252)	(0.267)	(0.230)
White Females	3291	699						
1st Stage				-0.000146**	0.051**	0.042**	0.027**	
				(0.000051)	(0.021)	(0.014)	(0.012)	0.01
Estimated Coefficients			0.101**	-0.373	-0.083	-0.117	-0.395	-0.001
			(0.018)	(0.341)	(0.328)	(0.288)	(0.437)	(0.231)
Black Males	708	286						
1st Stage				-0.000077	0.092*	0.027	0.036	
				(0.000136)	(0.055)	(0.037)	(0.033)	0.34
Estimated Coefficients			0.137**	0.824	-0.456	0.249	0.545	-0.141
			(0.035)	(1.741)	(0.639)	(1.053)	(0.706)	(0.386)
Black Females	764	288						
1st Stage				-0.00012	-0.037	0.007	0.001	
				(0.00012)	(0.046)	(0.035)	(0.029)	0.65
Estimated Coefficients			0.051*	0.017	-0.863	5.224	3.464	0.205
			(0.031)	(0.773)	(1.419)	(25.465)	(78.730)	(0.463)

Notes: First stage results are from a regression of the binary variable for repeating a grade (1=repeat) on the instrument(s) and the set of control variables. 2nd stage and OLS results use the binary variable for dropping out of high school (1=dropout) as the dependent variable. Standard errors in parenthesis (corrected for heteroskedasticity and clustering). **denotes significance at the 5% level for 2-tailed test. Numbers in the first stage section of column 8 are the p-values for the test that the month dummies jointly equal zero. Additional variables included in all regressions: indicator variables of parent's educational attainment, family income indicators, indicators for urban and rural status of students location in base sample year, and census region (4 sector classification).

BEST COPY AVAILABLE

Table 5

Estimates of the Effect of Grade Retention on Average Log Earnings

			OLS Estimates	Instrumental Variable Estimates				
					Binary Measures			
	Sample Size	*# of Schools*		*ndays*	*ndays<=30*	*ndays<=60*	*ndays<=90*	*Month Indicators*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
White Males	2551	673						
1st Stage				-0.000206**	0.075**	0.064**	0.058**	
				(0.000067)	(0.029)	(0.021)	(0.017)	0.02
Estimated Coefficients			-0.092**	0.230	0.310	0.228	0.316	0.178
			(0.039)	(0.653)	(0.741)	(0.625)	(0.558)	(0.538)
Black Males	477	239						
1st Stage				-0.000134	0.070	0.039	0.037	
				(0.000165)	(0.063)	(0.047)	(0.040)	0.60
Estimated Coefficients			-0.152	3.013	3.696	2.377	2.998	2.377
			(0.112)	(4.708)	(3.842)	(4.105)	(4.093)	(1.969)

Notes: First stage results are from a regression of the binary variable for repeating a grade (1=repeat) on the instrument(s) and the set of control variables. 2nd stage and OLS results use the average of log real earnings for 1990-1992 as the dependent variable. Standard errors in parenthesis (corrected for heteroskedasticity and clustering). **denotes significance at the 5% level for 2-tailed test. Numbers in the first stage section of column 8 are the p-values for the test that the month dummies jointly equal zero. Additional variables included in all regressions: indicator variables of parent's educational attainment, family income indicators, indicators for urban and rural status of students location in base sample year, and census region (4 sector classification).

BEST COPY AVAILABLE

Table A1
Kindergarten Entrance Dates

Date	States
Jan 1	CT, DE
Aug 1	MO
Aug 31	ND, WA
Sep 1	AZ, FL, GA, KS, MA, MN, MS, NM, OK, OR, SD, TX, UT, WI, WV
Sep 10	MT
Sep 15	IA, WY
Sep 30	NV, OH, TN, VA
Oct 1	AL, AR, KY
Oct 15	ID, ME, NE
Oct 16	NC
Nov 1	IL, SC
Nov 2	AK
Dec 1	CA, MI, NY
Dec 31	DC, HI, MD, RI
Local Discretion	CO, IN, LA, NH, NJ, PA, VT

Source: Wolf and Kessler (1987), Table 2

Table A2
Descriptive Statistics

N=7809

	Mean	Std. Dev.	Min	Max
High School Dropout	0.16	0.37	0	1
Ever Retained	0.14	0.35	0	1
White	0.81	0.39	0	1
Black	0.19	0.39	0	1
Male	0.48	0.50	0	1
College Graduate, Father	0.21	0.41	0	1
College Graduate, Mother	0.14	0.35	0	1
Missing Father's Education	0.13	0.34	0	1
Missing Mother's Education	0.09	0.28	0	1
Income Categories				
8000-14999	0.11	0.31	0	1
15000-19999	0.10	0.30	0	1
20000-24999	0.11	0.31	0	1
25000-29999	0.11	0.31	0	1
30000-39999	0.13	0.34	0	1
40000-49999	0.07	0.26	0	1
50000 +	0.09	0.29	0	1
Missing Income	0.24	0.43	0	1
High School Information				
urban	0.20	0.40	0	1
rural	0.29	0.45	0	1
West Region	0.15	0.36	0	1
North East Region	0.16	0.37	0	1
Central Region	0.35	0.48	0	1
Instruments				
ndays	182	106	0	365
ndays<=30	0.08	0.28	0	1
ndays<=60	0.17	0.38	0	1
ndays<=90	0.26	0.44	0	1